# 2010~2020 中国内地小学生心理健康问题检出率的元分析*

黄潇潇[1] 张亚利[1] 俞国良[2]

([1]中国人民大学教育学院; [2]中国人民大学心理研究所, 北京 100872)

**摘 要** 我国小学生心理健康问题的检出率莫衷一是，为整合调查结果并分析其原因，对2010~2020 年来检索到的相关研究进行了元分析。结果显示：(1)小学生心理健康问题的检出率由高到低依次是睡眠问题(25.2%, 95%CI = [0.160, 0.373])、抑郁(14.6%, 95%CI = [0.120, 0.175])、焦虑(12.3%, 95%CI = [0.063, 0.227])、攻击行为(4.1%, 95%CI = [0.017, 0.099])、退缩(3.8%, 95%CI = [0.023, 0.062])、违纪行为(3.7%, 95%CI = [0.018, 0.073])、躯体化(3.6%, 95%CI = [0.017, 0.073])。(2)测量工具、检出标准、检出时间是导致小学生心理健康问题检出率不一致的关键因素。总体而言，我国小学生除了睡眠问题、抑郁和焦虑检出率偏高外，整体心理健康状况尚可。未来应开发适合中国小学生的心理健康状况测评工具及检出标准，为心理健康问题的预防和精准干预提供支持。

**关键词** 小学生；心理健康问题；检出率；元分析

## 1 引言

随着我国社会经济的快速发展，竞争日益加剧，使得人们常常在人际交往、社会适应等方面产生心理困扰。近来研究发现，心理健康问题已逐渐从成人波及到青少年、儿童，呈现“低龄化”趋势(杜柏玲，2016；余欣欣等，2019)。据 2020 年世界卫生组织估计，个体约 50%的心理问题出现在 14 岁前，即小学阶段。小学生正处于身心快速发展之际，是自我意识、人格成长、个性形成的关键时期，其心理健康状况严重影响着学业成就与个体幸福感。因而，若小学生心理健康问题未能得到及时筛查、预防和干预，有可能发展为终生心理疾病，难以矫正治疗。为此，作为改善与提升小学生心理健康的基本前提，摸清小学生心理健康问题检出率，从而了解小学生心理健康真实状况，以便日后针对性的检测和干预就显得尤为重要。

尽管目前有关小学生心理健康问题检出率的文章并不少见，但报告的检出率因研究设计、测量工具、地区和样本量的不同而有很大差异。如赵玫(2010)调查出有异常心理问题

---

收稿日期：2021-08-25

* 中国人民大学“双一流”跨学科（心理健康教育）重大创新规划平台支持。

通信作者: 俞国良, E-mail: yugllxl@sina.com

倾向的小学生占 25.63%，而张柳和荆玉梅(2019)则发现其心理健康问题检出率仅为 7.9%。因而我国小学生心理健康问题检出率仍然众说纷纭，未有定论。近来，元分析方法的出现和成熟为解决该问题提供了可能。它通过定量手段整合已有研究，不仅可以判别小学生心理健康问题的数量、检出比例和分布特征等，还能整体把握小学生心理健康总体动向和趋势。近年来，已有个别研究利用元分析估计了小学生个别心理健康问题的检出率，如 Li 等(2019)发现小学生抑郁症状检出率为 17.5%。然而，既有元分析仍存在不足之处。首先，仅揭示了部分心理健康问题(如抑郁、焦虑)的检出率，并没有涵盖更为全面的心理健康问题；其次，未针对性地分析近 10 年来的心理健康问题。快速的社会变化是影响心理健康水平的重要因素，近 10 年来我国经济、文化等各领域都产生了巨大变迁，因而小学生心理健康状况也可能发生变化。最后，对调节变量的纳入和探讨不够具体。因此，立足于近 10 年来研究成果，采用元分析方法整合小学生心理健康问题检出率并探讨其影响因素是十分必要的。

综上，本研究主要目标是：(1)利用元分析总结 2010~2020 年有关小学生心理健康问题的检出率。根据皮亚杰的认知发展理论，小学生处于具体运算阶段，其情感、社会性发展、认知发展三者并行发展，都具有从“自我中心”向“脱自我中心”转移的特点(刘国雄，张丽锦，2010)，即将他人作为认知、情感对象。这也就表明，小学生的情感复杂性以及社会化程度都将随着步入学校生活而开始飞速提高。因而，小学生的心理健康问题尤其是情绪情感问题、社会行为问题应受到广泛关注。由于涉及到的心理健康问题繁多，无法覆盖全部指标，故本研究主要借鉴 Achenbach(1966)提出的儿童青少年心理病理性问题两分法，关注内化问题(包括焦虑、抑郁、躯体化、退缩、睡眠问题)和外化问题(包括攻击行为、违纪行为)。(2)考察调节变量对小学生心理健康问题检出率的影响。具体来说，将从出版年代、测量工具、检出时间、性别、区域、年级、独生与否、生源地等方面讨论调节变量的作用，试图全面展现小学生心理健康问题检出率状况。

## 2 方法

### 2.1 文献检索与筛选

本研究中所包括的心理健康问题指标较多，因而检索过程是根据不同心理健康问题独立开展的。与既有类似研究不同的是，为了更加全面的纳入文献，避免纳入偏倚的出现，本研究还尽可能地检索了未发表的文献。另外，在检索时对于关心的特定学生群体也未加限定，而是在全面检索之后，再进行筛选。中文使用的是中国知网期刊全文数据库、硕博论文数据库，英文使用的是 Web of Science 核心合集数据库，检索范围设定为摘要。在中文

数据库检索时，焦虑需包含关键词“焦虑”和“检出率”；抑郁需包含关键词“抑郁”和“检出率”；躯体化需包括关键词“躯体化”或“躯体主诉”和“检出率”；退缩需包括关键词“退缩”和“检出率”；睡眠问题需包括关键词“睡眠问题”或“睡眠障碍”或“失眠”和“检出率”；攻击行为需包括关键词“攻击”或“欺负”和“检出率”；违纪行为需包括关键词“违纪”和“检出率”。在英文数据库检索时，焦虑需包含关键词“anxi*” and “prevalence” or “detection rate” and “Chin*”；抑郁需包含关键词“depress*” and “prevalence” or “detection rate” and “Chin*”；躯体化需包括关键词“somati*” and “prevalence” or “detection rate” and “Chin*”；退缩需包括关键词“withdrawal” and “prevalence” or “detection rate” and “Chin*”；睡眠问题需包括关键词“sleep” or “insomnia” and “prevalence” or “detection rate” and “Chin*”；攻击行为需包括关键词“aggress*” or “bully*” and “prevalence” or “detection rate” and “Chin*”；违纪行为需包括关键词“disciplin*” and “prevalence” or “detection rate” and “Chin*”。检索时间段为 2010 年 1 月 1 日至 2020 年 12 月 31 日。共获取文献：焦虑 3126 篇，抑郁 3844 篇，躯体化 551 篇，退缩 167 篇，睡眠问题 1255 篇，攻击行为 469 篇，违纪行为 126 篇。

对于搜索到的文献，利用 EndNote X9 软件导入，按照如下标准筛选：(1)文献需为实证研究且为一手资料，纯理论及综述性研究将被排除。(2)报告了检出率或能够计算检出率的必要信息，对数据不清晰或存在明显错误且无法进行修正的研究予以删除。(3)必须对测量工具有明确介绍。(4)同一批数据重复发表的仅取其一。(5)研究对象为中国小学阶段或对应年龄段(6~12 岁)学生。(6)研究对象非特殊人群，如学困小学生、听力障碍儿童等。文献筛选过程先由 1 位博士生完成，最后按心理健康指标分配给三位博士生随机抽取 10%对剔除文献和保留文献进行核查，出现不一致结果需要与最初文献筛选者协商解决。最终纳入的文献包括：焦虑 9 篇，抑郁 34 篇，躯体化 12 篇，退缩 10 篇，睡眠问题 11 篇，攻击行为 13 篇，违纪行为 12 篇。文献纳入流程如下图 1 所示。

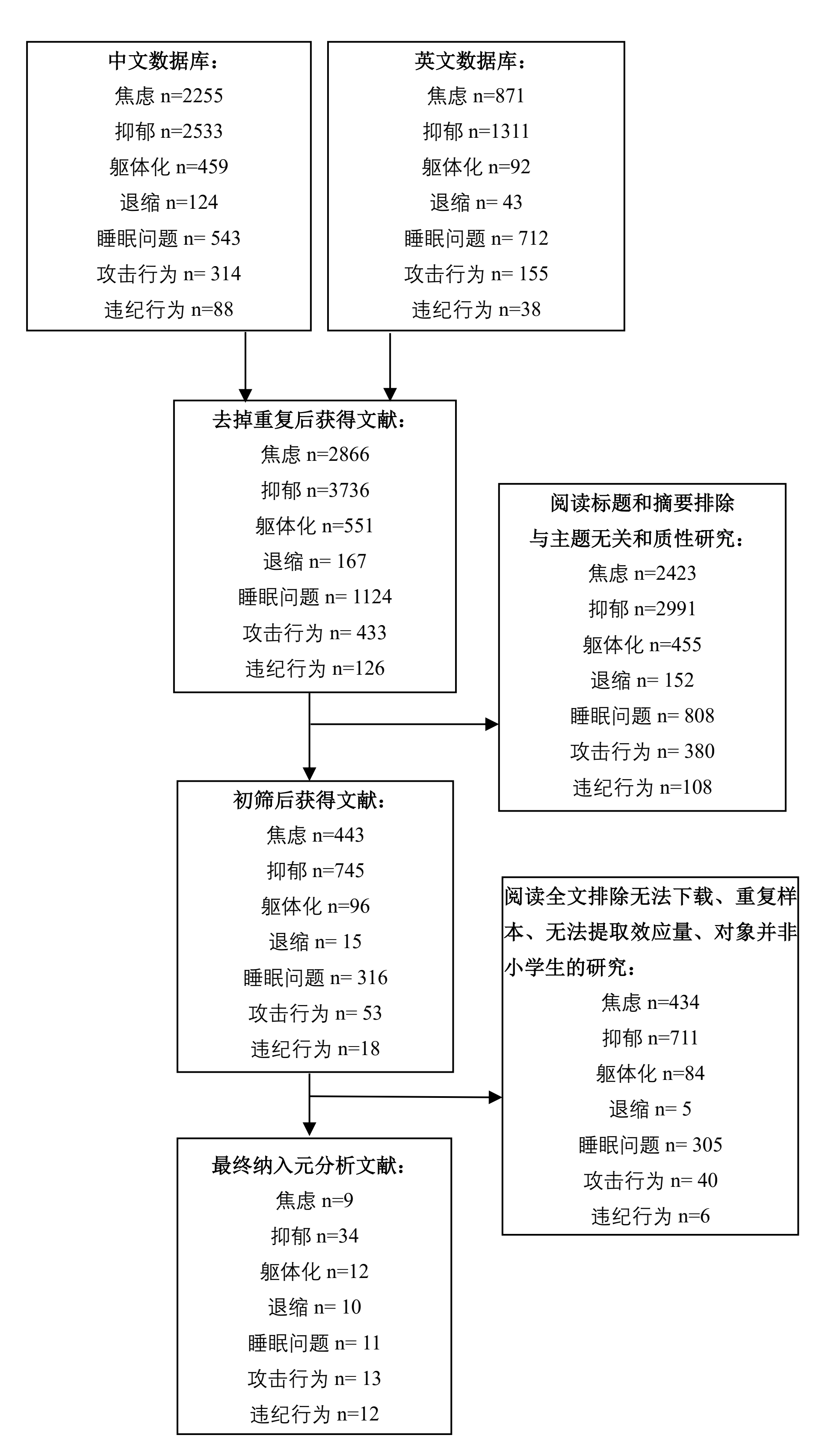



**图 1 文献纳入流程**

## 2.2 文献编码

对纳入的每项研究根据以下特征进行编码：文献名、第一作者、发表年份、是否发表、

取样地区、被试数、被试平均年龄、测量工具、检出标准、检出时间、检出率。除了录入总检出率外，还将其它分组报告结果(包括性别、年龄、生源地等)录入到该项研究的子数据库中。文献编码先由一位博士生完成，完成后再由另一位博士生对照原文核查。若核查出与原文不一致的情况，经过讨论后将对数据进行更正。最终纳入文献的详细信息均可开放获取(https://osf.io/7uewg/)。

### 2.3 出版偏倚控制与检验

出版偏倚是指具有统计学意义的研究结果比无统计学意义的结果更容易出版(Kuppens et al., 2013)。因此，仅检索已经发表的文献将无法代表整个领域的研究现状。为了尽量避免出版偏倚问题，本研究除了纳入发表文献外，还检索了未公开发表的硕博论文。此外，本研究还将采用漏斗图(Funnel plot)和 Kendall 等级相关检验检验是否存在出版偏倚。于漏斗图而言，由于样本数量增加会导致数据的随机变化减小，因此，如果研究不存在明显的发表偏倚，则图形将呈现一个倒漏斗状对称分布；于 Kendall 等级相关检验而言，其 Tau 值(Kendall’s tau) 表示效应值与其标准差的相关程度，不显著则表明出版偏倚较小。

### 2.4 模型选择

目前，在元分析中，估计总体效应值的方法主要有固定效应模型和随机效应模型，其区别是代表的权重成分不同。具体而言，固定效应模型假设，在所有原始研究中得出的测量结果由相同的真值和随机误差两部分组成，而每个效应量的不同则全部是由抽样误差造成。相反，随机效应模型则假设，研究测量的结果由真值、系统误差和抽样误差三部分组成，每个测量结果的差异除了受抽样误差影响之外，还受不同种类的系统误差影响，比如测量工具、检出标准等。从文献编码过程中可以发现，研究中所包含分组报告的结果不尽相同，因而此时选择随机效应模型将比固定效应模型更为合理。除此以外，本研究还将通过 $Q$ 检验结果以及 $I^2$ 值两个指标进行异质性检验，以查看该模型选择的适切性。根据 Higgins 等(2003)的观点，如果 $Q$ 检验结果显著或 $I^2$ 的值高于 75%，说明在本研究中，相较固定效应模型，选择随机效应模型更加合适。

### 2.5 数据处理

本研究采用 CMA3.3 软件(Comprehensive Meta-Analysis Version 3.3)进行小学生心理健康问题检出率的元分析，包括总体检出率估计和调节变量的检验。计算合并检出率时，CMA 软件会将输入的比率数据转化为 logit 数据[logit=Log(p/(1- p))]进行分析(var(logit) = $\frac{1}{case}+\frac{1}{non_case}$)，并在分析完成后将 logit 结果再次转换为比率数据输出(Card, 2012)。调节变量的分析主要采用两种形式：(1)当调节变量为连续变量时，采用元回归分析考察结果是否

显著；(2)当调节变量为分类变量时，采用亚组分析检验结果是否显著。

# 3 结果

## 3.1 文献纳入情况

本研究中，(1)焦虑共纳入研究 9 篇(9 个效应值，12691 名被试)，均为期刊论文；(2)抑郁共纳入研究 34 篇(34 个效应值，94926 名被试)，包括硕博论文 8 篇，期刊论文 26 篇；(3)躯体化共纳入研究 12 篇(12 个效应值，41588 名被试)，包括硕博论文 3 篇，期刊论文 9 篇；(4)退缩共纳入研究 10 篇(10 个效应值，37533 名被试)，包括硕博论文 4 篇，期刊论文 6 篇；(5)睡眠问题共纳入研究 11 篇(11 个效应值，17437 名被试)，包括硕博论文 1 篇，期刊论文 10 篇；(6)攻击行为共纳入研究 13 篇(13 个效应值，46042 名被试)，包括硕博论文 6 篇，期刊论文 7 篇；(7)违纪行为共纳入研究 12 篇(26 个效应值，39159 名被试)，包括硕博论文 3 篇，期刊论文 9 篇。

## 3.2 异质性检验

异质性检验结果见表 1。各心理健康问题指标的 $Q$ 值均显著，且 $I^2$ 值均超过了 75%的标准，表明随机效应模型的选取是恰当的。该结果也提示不同研究间的估计值差异可能受到了一些研究特征因素的干扰，可进行调节变量的效应分析。

## 3.3 主效应检验

采用随机效应模型，分别对纳入的小学生心理健康问题检出率进行分析，结果显示：(1)焦虑的检出率为 12.3%；(2)抑郁的检出率为 14.6%；(3)躯体化的检出率为 3.6%；(4)退缩的检出率为 3.8%；(5)睡眠问题的检出率为 25.2%；(6)攻击行为的检出率为 4.1%；(7)违纪行为的检出率为 3.7%(表 1)。

**表 1 小学生心理健康问题的检出率**

| 心理健康问题种类 | k | 被试数 | 效应值及 95%的置信区间 | | | 异质性检验 | |
|---|---|---|---|---|---|---|---|
| | | | 检出率 | 下限 | 上限 | $I^2$ | $p$ |
| 焦虑 | 9 | 12691 | 0.123 | 0.063 | 0.227 | 99.149 | ＜0.001 |
| 抑郁 | 34 | 94926 | 0.146 | 0.120 | 0.175 | 99.154 | ＜0.001 |
| 躯体化 | 12 | 41588 | 0.036 | 0.017 | 0.073 | 99.412 | ＜0.001 |
| 退缩 | 10 | 37553 | 0.038 | 0.023 | 0.062 | 98.642 | ＜0.001 |
| 睡眠问题 | 11 | 17437 | 0.252 | 0.160 | 0.373 | 99.505 | ＜0.001 |
| 攻击行为 | 13 | 46042 | 0.041 | 0.017 | 0.099 | 99.728 | ＜0.001 |
| 违纪行为 | 12 | 39159 | 0.037 | 0.018 | 0.073 | 99.310 | ＜0.001 |

为说明主效应检验结果的稳健性，敏感性分析发现，在排除任意一个样本后，焦虑的总体检出率在 11.1%～15.1%之间浮动；抑郁的总体检出率在 13.3%~ 14.6%之间浮动；躯体化的总体检出率在 2.9% ～ 4.1%之间浮动；退缩的总体检出率在 3.2% ～ 4.2%之间浮动；睡眠问题的总体检出率在 21.7%~29.1%之间浮动；攻击行为的总体检出率在 3.6%~ 4.7%之间

浮动，违纪行为的总体检出率在 3.0%~ 4.2%之间浮动，均与各自的总体估计值相差不大，说明元分析估计结果具有较高稳定性。

### 3.4 调节效应检验

就焦虑而言，如表 2 所示，调节效应检验发现：(1)出版年代的调节效应不显著($b=0.04$, 95%CI=[-0.12, 0.21])，说明近 10 年来焦虑问题的检出率无显著变化；(2)测量工具和标准的调节效应不显著，虽然使用不同的工具及筛查标准其焦虑检出率有所差别，但差异并不显著。(3)检出时间的调节作用显著，3 个月的检出率显著大于 1 周的检出率，表明检出时间可能是焦虑检出率高低的重要因素。(4)区域的调节作用不显著。东部、中部、西部的焦虑检出率无明显差异。

**表 2 焦虑检出率的调节效应分析**

| 调节变量 | 异质性检验 | | | 类别 | $k$ | 检出率 | 95% CI | |
|---|---|---|---|---|---|---|---|---|
| | $Q_B$ | $df$ | $p$ | | | | 下限 | 上限 |
| 测量工具+检出标准 | 2.669 | 2 | 0.263 | Conners | 2 | 0.085 | 0.021 | 0.293 |
| | | | | SCARED23 | 2 | 0.146 | 0.037 | 0.432 |
| | | | | SCARED22 | 2 | 0.345 | 0.106 | 0.701 |
| 检出时间 | 11.986 | 1 | 0.001 | 1 周 | 2 | 0.038 | 0.015 | 0.093 |
| | | | | 3 个月 | 4 | 0.231 | 0.138 | 0.361 |
| 区域 | 0.169 | 2 | 0.919 | 东北 | - | - | - | - |
| | | | | 东部 | 2 | 0.077 | 0.012 | 0.255 |
| | | | | 西部 | 3 | 0.110 | 0.027 | 0.360 |
| | | | | 中部 | 3 | 0.120 | 0.029 | 0.383 |

注：Conners 为 Conners 父母症状问卷(Conners, 1969)；SCARED23 为儿童焦虑性情绪障碍筛查表(Birmaher et al., 1997)；SCARED22 修订为陈洁琼(2012)修订的儿童焦虑性情绪障碍筛查表。工具后的数字为检出标准，下同。

就抑郁而言，如表 3 所示：(1)出版年代的调节效应不显著($b=0.01$, 95%CI=[-0.07, 0.10])，说明近 10 年来抑郁检出率虽然有上升趋势，但变化不明显；(2)测量工具的调节效应显著。不同工具和标准测得检出率在 5.7%~17.4%之间，CBCL 测得的检出率最低，CESD 测得的检出率最高。(3)检出时间的调节效应显著。6 个月的检出率最低，1 周的检出率最高，此倒挂现象可能与采用的测量工具有关。(4)性别的调节作用不显著，仅呈现男高女低的倾向。(5)独生与否的调节作用不显著。(6)年级的调节作用不显著，但四年级至六年级的检出率随年级增加呈现递减趋势。(7)生源地和区域的调节作用不显著。

**表 3 抑郁检出率的调节效应分析**

| 调节变量 | 异质性检验 | | | 类别 | $k$ | 检出率 | 95% CI | |
|---|---|---|---|---|---|---|---|---|
| | $Q_B$ | $df$ | $p$ | | | | 下限 | 上限 |
| 测量工具 | 16.458 | 4 | 0.002 | CBCL | 5 | 0.057 | 0.034 | 0.094 |
| | | | | CDI | 16 | 0.165 | 0.130 | 0.206 |
| | | | | CESD | 6 | 0.174 | 0.119 | 0.249 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | DSRSC | 4 | 0.160 | 0.099 | 0.250 |
| | | | | MFQ | 2 | 0.171 | 0.086 | 0.309 |
| 检出时间 | 9.702 | 2 | 0.008 | 1 周 | 13 | 0.192 | 0.146 | 0.249 |
| | | | | 2 周 | 15 | 0.147 | 0.112 | 0.190 |
| | | | | 6 个月 | 5 | 0.057 | 0.033 | 0.097 |
| 性别 | 0.389 | 1 | 0.533 | 男生 | 17 | 0.109 | 0.082 | 0.144 |
| | | | | 女生 | 17 | 0.096 | 0.096 | 0.127 |
| 独生与否 | 2.509 | 1 | 0.113 | 独生 | 3 | 0.117 | 0.075 | 0.177 |
| | | | | 非独生 | 3 | 0.187 | 0.124 | 0.271 |
| 年级 | 0.030 | 2 | 0.985 | 四年级 | 6 | 0.194 | 0.130 | 0.279 |
| | | | | 五年级 | 6 | 0.187 | 0.124 | 0.272 |
| | | | | 六年级 | 7 | 0.186 | 0.128 | 0.262 |
| 生源地 | 0.012 | 1 | 0.913 | 城镇 | 5 | 0.082 | 0.023 | 0.255 |
| | | | | 农村 | 5 | 0.090 | 0.025 | 0.276 |
| 区域 | 3.705 | 2 | 0.157 | 东北 | - | - | - | - |
| | | | | 东部 | 9 | 0.108 | 0.069 | 0.167 |
| | | | | 西部 | 10 | 0.181 | 0.121 | 0.263 |
| | | | | 中部 | 11 | 0.114 | 0.071 | 0.170 |

注：CBCL 量表为 Achenbach 儿童行为量表(Achenbach & Edelbrock, 1979)；CDI 为儿童抑郁量表(Kovacs, 1992)；CESD 为流调中心抑郁量表(Radloff, 1997)；DSRSC 为儿童抑郁自评量表(Birleson, 1981)；MFQ 为情绪问卷(Wood, et al., 1995)。

就躯体化而言，如表 4 所示：(1)出版年代的调节效应不显著($b$ = -0.04, 95%CI=[-0.23, 0.15])，说明近 10 年来躯体化的检出率虽有下降趋势，但变化不显著；(2)检出标准的调节作用显著。依据中国常模标准化的 CBCL6 量表，其躯体化检出率显著大于 CBCL98 量表。(3)虽然男生的躯体化检出率稍高，但性别的调节作用不显著。(4)区域的调节效应不显著，表明东部、中部、西部地区的躯体化检出率无显著差异。

**表 4 躯体化检出率的调节效应分析**

| 调节变量 | 异质性检验 | | | 类别 | $k$ | 检出率 | 95% CI | |
|---|---|---|---|---|---|---|---|---|
| | $Q_B$ | $df$ | $p$ | | | | 下限 | 上限 |
| 测量工具+检出标准 | 12.898 | 1 | <0.001 | CBCL98 | 6 | 0.015 | 0.009 | 0.026 |
| | | | | CBCL6 | 5 | 0.063 | 0.036 | 0.108 |
| 性别 | 0.022 | 1 | 0.883 | 男生 | 8 | 0.033 | 0.012 | 0.089 |
| | | | | 女生 | 8 | 0.029 | 0.010 | 0.080 |
| 区域 | 5.076 | 2 | 0.079 | 东北 | - | - | - | - |
| | | | | 东部 | 6 | 0.025 | 0.012 | 0.052 |
| | | | | 西部 | 2 | 0.045 | 0.012 | 0.152 |
| | | | | 中部 | 3 | 0.107 | 0.038 | 0.269 |

注：CBCL98 代表该量表分数超过第 98 百分位即被认为异常；CBCL6 为忻仁娥与张志雄(1992)根据全国 22 个省市的青少年儿童所作的 Achenbach 儿童行为量表标准化分类，超过 6 分则为异常，下同。

就退缩而言，如表 5 所示，调节效应检验发现：(1)出版年代的调节效应不显著($b$ = 0.008, 95%CI=[-0.13, 0.14])，说明近 10 年来检出率虽有上升趋势，但变化不明显；(2)检出标准的调节效应不显著。

**表 5 退缩检出率的调节效应分析**

| 调节变量 | 异质性检验 | | | 类别 | $k$ | 检出率 | 95% CI | |
|---|---|---|---|---|---|---|---|---|
| | $Q_B$ | $df$ | $p$ | | | | 下限 | 上限 |
| 测量工具+检出标准 | 1.387 | 1 | 0.239 | CBCL98 | 6 | 0.030 | 0.016 | 0.056 |
| | | | | CBCL6 | 4 | 0.055 | 0.025 | 0.114 |

就睡眠问题而言，如表 6 所示：(1)出版年代的调节效应不显著($b$ = -0.13, 95%CI=[-0.27, 0.01])，说明近 10 年来虽然睡眠问题检出率随年份有下降趋势，但变化不显著。(2)测量工具的调节效应不显著，CSHQ 的检出率虽然大于 PSQI，但两者之间的差别并不明显。(3)检出时间的调节效应不显著。(4)性别的调节效应不显著，仅呈现男高女低的倾向。(5)区域的调节效应不显著。

**表 6 睡眠问题检出率的调节效应分析**

| 调节变量 | 异质性检验 | | | 类别 | $k$ | 检出率 | 95% CI | |
|---|---|---|---|---|---|---|---|---|
| | $Q_B$ | $df$ | $p$ | | | | 下限 | 上限 |
| 测量工具 | 2.922 | 1 | 0.087 | PSQI | 3 | 0.208 | 0.087 | 0.418 |
| | | | | CSHQ | 3 | 0.476 | 0.250 | 0.713 |
| 检出时间 | 4.92 | 2 | 0.086 | 过去四周 | 3 | 0.476 | 0.278 | 0.682 |
| | | | | 近一个月 | 5 | 0.211 | 0.121 | 0.342 |
| | | | | 无 | 2 | 0.268 | 0.114 | 0.511 |
| 性别 | 0.949 | 1 | 0.330 | 男生 | 9 | 0.041 | 0.017 | 0.096 |
| | | | | 女生 | 9 | 0.032 | 0.013 | 0.077 |
| 区域 | 0.955 | 2 | 0.620 | 东北 | - | - | - | - |
| | | | | 东部 | 5 | 0.0260 | 0.134 | 0.443 |
| | | | | 西部 | 3 | 0.0266 | 0.113 | 0.509 |
| | | | | 中部 | 2 | 0.0422 | 0.167 | 0.727 |

注：PSQI 为匹兹堡睡眠指数量表(Buysse et al., 1989)；CSHQ 为儿童睡眠习惯问卷(Owens et al., 2000)。

就攻击行为而言，如表 7 所示：(1)出版年代的调节效应不显著($b$ = 0.06, 95%CI=[-0.10, 0.23])，说明近 10 年来小学生攻击行为的检出率虽然略有上升趋势，但变化并不明显。(2)测量工具和标准的调节效应显著，BWAQ 的检出率最高，CBCL6 次之，检出率最低的则是 CBCL98。提示测量工具和标准是影响小学生攻击行为检出率的重要调节变量。(3)检出时间的调节效应显著。未写明检出时间的检出率显著高于最近半年的检出率，说明检出时间的长短很可能影响着小学生攻击行为检出率的高低。(4)尽管男生有高于女生的趋势，但性别的调节效应仍然不显著。(5)区域的调节效应显著，中部地区小学生攻击行为的检出率明显高于东部地区。

**表 7 攻击行为检出率的调节效应分析**

| 调节变量 | 异质性检验 | | | 类别 | $k$ | 检出率 | 95% CI | |
|---|---|---|---|---|---|---|---|---|
| | $Q_B$ | $df$ | $p$ | | | | 下限 | 上限 |
| 测量工具+检出标准 | 41.611 | 2 | <0.001 | BWAQ | 2 | 0.196 | 0.110 | 0.324 |
| | | | | CBCL98 | 6 | 0.018 | 0.012 | 0.027 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | CBCL6 | 4 | 0.048 | 0.030 | 0.076 |
| 检出时间 | 22.878 | 1 | <0.001 | 无 | 3 | 0.188 | 0.098 | 0.331 |
| | | | | 最近半年 | 10 | 0.027 | 0.018 | 0.040 |
| 性别 | 0.039 | 1 | 0.865 | 男生 | 9 | 0.045 | 0.014 | 0.132 |
| | | | | 女生 | 9 | 0.039 | 0.012 | 0.117 |
| 区域 | 4.981 | 1 | 0.026 | 东北 | - | - | - | - |
| | | | | 东部 | 7 | 0.028 | 0.017 | 0.045 |
| | | | | 西部 | - | - | - | - |
| | | | | 中部 | 3 | 0.079 | 0.036 | 0.163 |

注：BWAQ 为 Buss-Warren 攻击问卷(Buss & Warren, 2000)。

就违纪行为而言，如表 8 所示：(1)出版年代的调节效应不显著($b$ = 0.08, 95%CI=[-009, 0.25])，说明近 10 年来违纪行为的检出率虽有上升趋势，但并不明显；(2)检出标准的调节效应显著，经过中国常模标准化的 CBCL6 量表，其违纪行为的检出率显著高于 CBCL98 量表。(3)性别调节作用不显著，仅可见男高女低的倾向。(4)区域的调节作用不显著。东部、西部和中部的小学生违纪行为检出率无明显差异。

**表 8 违纪行为检出率的调节效应分析**

| 调节变量 | 异质性检验 | | | 类别 | $k$ | 检出率 | 95% CI | |
|---|---|---|---|---|---|---|---|---|
| | $Q_B$ | $df$ | $p$ | | | | 下限 | 上限 |
| 测量工具+检出标准 | 4.434 | 1 | 0.035 | CBCL98 | 8 | 0.022 | 0.013 | 0.037 |
| | | | | CBCL6 | 3 | 0.063 | 0.028 | 0.136 |
| 性别 | 0.144 | 1 | 0.704 | 男生 | 9 | 0.041 | 0.017 | 0.096 |
| | | | | 女生 | 9 | 0.032 | 0.013 | 0.077 |
| 区域 | 0.679 | 2 | 0.712 | 东北 | - | - | - | - |
| | | | | 东部 | 7 | 0.035 | 0.013 | 0.096 |
| | | | | 西部 | 2 | 0.078 | 0.012 | 0.377 |
| | | | | 中部 | 2 | 0.027 | 0.003 | 0.184 |

### 3.5 发表偏差检验

首先，就漏斗图而言，小学生各类心理健康问题的效应值均聚集在漏斗图上端，且基本分散于效应值平均值的两端，初步显示发表偏差问题并不明显。其次，就 Kendall 等级相关检验结果来看，焦虑(Tau = -0.08, $p$ = 0.75)、抑郁(Tau = -0.19, $p$ = 0.11)、躯体化(Tau =-0.25, $p$ = 0.24)、退缩(Tau = 0.09, $p$ = 0.72)、睡眠问题(Tau = 0.11, $p$ = 0.64)、攻击行为(Tau = 0.14, $p$ = 0.50)、违纪行为(Tau = -0.02, $p$ = 0.95)的结果均不显著，说明本研究不存在明显的发表偏差。

## 4 讨论

本研究利用元分析方法，囊括了较为全面的心理健康问题指标，对近 10 年来(2010~2020)小学生心理健康问题的检出率进行了定量整合，并考察了背后的影响因素，在一定程度上澄清了我国近 10 年来(2010~2020)小学生各类心理健康问题的整体检出率。

## 4.1 小学生心理健康问题的检出率

研究结果显示，小学生焦虑检出率为 12.3%，抑郁检出率为 14.6%，躯体化检出率为 3.6%，退缩检出率为 3.8%，睡眠问题检出率为 25.2%，攻击行为检出率为 4.1%，违纪行为检出率为 3.7%。从上述结果可见，对于小学生而言，不同心理健康问题的检出率差异较大，如躯体化检出率低至 3.6%，而睡眠问题检出率则高至 25.2%。该结果表明，小学生的心理健康问题不容忽视，尤其是检出率较高的前三位——睡眠问题、抑郁和焦虑。

第一是睡眠问题。本研究由于纳入了未发表文献，且考察了近 10 年研究成果，使得检出率与 Chen 等(2021) 元分析中 38.9%的结果不同，但两者都说明我国小学生睡眠问题已较为普遍。这不仅是因为很多睡眠疾病(如夜惊、噩梦、入睡困难等)在儿童期更为高发(张斌等，2013)，同时也与我国课业负担较重，难以保证高质量、长时间的睡眠活动有关(龙鑫等，2020)。此外，小学生电子产品过度使用、手机成瘾现象较为严重(来枭雄等，2020)，容易导致入睡困难等现象发生，需要引起社会各界对此的广泛重视。第二是抑郁。本结果与《心理健康蓝皮书》中约一成的小学生抑郁检出率结果相似(傅小兰等，2021)，但却低于 Li 等(2019)元分析中得出的中国小学生 17.5%的检出率，也低于刘福荣等(2021)元分析 17.1%的结果，这可能是由于 Li 等(2019)中未包含 200 人以下的样本数量，而刘福荣等(2021)则未检索英文数据库。因而，本研究的估计结果更加全面精确。第三是焦虑。本结果与李梦龙等(2019)的研究存在不同。李梦龙等(2019)曾纳入 15 项小学生社交焦虑状况文献，发现留守小学生社交焦虑检出率为 18.22%，非留守小学生则为 13.76%。这种结果差异可能是源于，本结果针对的是一般焦虑，而非单纯的社交焦虑。与国外其它国家相比，Erskine 等(2017)发现高收入地区的青少年儿童，其焦虑检出率为 18.27%，低收入和中等收入地区的检出率为 1.53%，意味着经济可能会对青少年儿童焦虑的检出率产生影响。作为发展中国家，我国小学生焦虑的检出率也处于中等偏高位置，需引起警惕。

其余几类小学生心理健康问题的检出率虽不高，均在 3.6%~4.1%之间浮动。但值得注意的是，某些心理健康问题一旦出现，很有可能会造成严重后果。如小学生的攻击和违纪行为，不仅不利于其个性、社会化和良好品德的顺利形成，还会造成学校氛围的恶化，给校园安全埋下无法预估的隐患。因而在心理健康问题的筛查、诊断与干预中，也须重点关注。总之，本研究中所得近 10 年来我国小学生各类心理健康问题的检出率结果，有利于掌握我国小学生心理健康的真实状况，帮助实现国家对小学生心理健康水平的宏观把控，同时也为改进心理健康教育、完善相关心理健康政策提供了一定依据和参考。

4.2 小学生心理健康问题的调节变量

第一是出版年代。结果表明，尽管大多数心理健康指标随年份增长略有上升，但这种趋势实际上并不显著。类似地，廖友国和连榕(2019)通过横断历史研究发现，1986~2017 近三十年国民心理健康总体呈现缓慢向好趋势。然而与之相反，针对具体学生群体，辛自强和张梅(2009)的元分析结果却表明，1992~2005 年间中学生心理健康水平稳中有降；高中生在 1990~2004 年间，心理健康水平也呈缓慢下降趋势(俞国良等，2016)。原因可能在于，即使相关心理健康保障和服务体系逐渐完善，资金投入和师资力量不断扩大(王燕等，2020)，但社会生活节奏加快、竞争愈发激烈、父母期望过高，使得学生群体(包括小学生)仍然面临着来自父母、社会的多种压力和挑战(黄文治，王志刚，2020)。因而，近 10 年来我国小学生心理健康检出率，并未明显表现出逐年向好态势。

第二是测量工具和标准。除了焦虑、退缩和睡眠问题不显著外，其它心理健康问题指标均显著，提示今后对于量表的选择上要慎重。值得注意的是，即使选择了同一量表，在使用上也可能因检出标准不同，导致检出率结果的显著差异。具体来看，CBCL 量表是否经过中国样本的常模标准化，对小学生的躯体化、攻击与违纪行为的检出率造成了影响。因而，在小学生心理健康问题的筛检工作中，测评量表及检出标准的正确选择、测量实践的规范化及透明化操作，对准确地评估小学生心理健康状况极其重要。以往某些心理健康问题指标，其检出率高低不一，很可能是由于测量工具和标准不一致，导致结果出现严重分歧。例如，刘福荣等(2021)的元分析发现，CES-D 的检出率高达 22.2%，而 BID 量表所得的抑郁检出率仅为 1.9%。为了实现心理健康问题检出率的可比性，未来也应开发更科学的测量工具，以便统一检出标准，便于相互对照。 此外，测量工具检出率差异较大，也可能是因为目前测量工具仍基于量表总分或症状数量的简单加和来诊断是否出现某一心理健康问题(刘宇等，2021)。然而不同测量工具往往涵盖的症状内容和数量有所差异，致使同一心理健康问题很难具有相同的症状表现，其检出率大小也会受其影响。

第三是检出时间。除了睡眠问题的检出时间无显著调节作用外，其余心理健康指标的检出时间，均存在显著差异。一般来说，短时间的检出时间测得的检出率，应比长时间或未规定检出时间测得的检出率低。值得注意的是，小学生抑郁检出时间的检出率，呈现出倒挂现象，即 1 周>2 周>6 个月。这说明检出率结果可能受工具影响更强，而非检出时间。许多研究都显示，不同量表调查条目、调查维度不同，评价方法、检出时间也各不一样，容易对检出率造成影响(刘福荣等，2021)。另外，睡眠问题检出时间的检出率差别不大。此类结果表明，目前有关检出率元分析测量工具不统一，检出时间不明晰，难以使检出率

结果在同一维度上与其它研究进行比较(Jaafari et al., in press)。因而，未来仍然需要探寻更适宜、更统一的测量工具，以求得出更为可靠、更稳定的检出率结果。

第四是性别。近 10 年来，我国小学生心理健康问题虽然表现出了男高女低的倾向，但性别差异并不显著。以往对于性别的调节效应，不同的研究取得了不一致的结论。杜柏玲(2016)发现，小学生心理健康问题的检出率存在显著的性别差异，而另一些研究则认为性别差异不显著(卫萍等，2007；杨洪猛等，2003)。这种矛盾情况，很有可能是受到取样条件、所在地区等不同因素的影响。本研究通过元分析技术，在一定程度上平衡了随机因素，得出我国小学生心理健康问题检出率并未出现性别分岭的结论。原因在于，一方面，研究发现无论是学业表现，还是未来成就与人际关系，父母对不同性别的子女所寄予的期望都是一样的(程琳，2010)；另一方面，随着性别平等教育的逐步推进，小学生在日常学习和生活中接触到的性别角色榜样等性别线索也越来越多样化(李超群等，2021)，日益增强的认知能力如推理、判断等都使其性别刻板印象开始逐渐减弱。这些都使得男女小学生在面对心理健康问题时，可能较少受到性别因素的影响。未来也应对此类调节变量的效应进行深入的探讨和分析。

第五是区域。研究结果显示，大部分指标在区域上无显著差别，仅有攻击行为存在显著区域差异，即中部地区高于东部。近 10 年来，相比较为发达的东部地区，我国中部地区经济发展仍比较落后，心理健康教育投入相对不足，学生获取心理健康教育的途径较少、资源较为有限(俞国良等, 2017)。如研究显示，全国心理健康教育课开设情况呈“倒U型”，中部地区开课率显著低于东部和西部(徐美贞，2005)，因此中部地区的小学生在面临心理困扰时，可能难以寻求及时而有效的帮助，导致心理健康问题检出率较高。该结论也得到了相关实证研究的支持，即经济水平较为低下的青少年儿童，更易产生敌对情绪，人际关系更为紧张，攻击行为的发生率可能将会更高(赵建华，2005)。

综上，从本研究中的调查结果可知，近 10 年来我国小学生心理健康总体水平较好，但部分心理健康问题如睡眠问题、抑郁、焦虑仍较为普遍，并且这些心理健康问题的检出率受到了测量工具、检出标准、检出时间等多因素影响。因此，重视小学生心理健康问题的筛检工作，加强小学生的心理健康教育十分必要。

### 4.3 研究不足与展望

第一，本研究纳入较多研究指标，因而在英文数据库中仅检索了 Web of Science 核心合集。虽然该平台文献较为全面权威，但受数据库影响，本研究可能会遗漏个别研究结果。未来研究应检索更多数据库，以求更完备地展现元分析数据。第二，本研究在开始前未进

行预注册，未来研究者应重视此项工作，通过衡量、印证预先假设和后期检验，从而进一步加强元分析研究的规范性、科学性和严谨性。第三，本研究主要借鉴内外化问题分析框架，关注小学生群体中较为典型的几类心理健康问题，未来的元分析可纳入更多指标，以求更全面地展现小学生不同心理健康问题的检出率状况。第四，本研究集中在现象调查，尚未关注到如何提高和改善小学生心理健康。未来的元分析研究可以重点整合提升心理健康的具体干预手段。

# 5 结论

(1)小学生面临着不同程度的心理健康问题，其中睡眠问题、抑郁和焦虑排在前三位。(2) 测量工具与检出标准是造成小学生心理健康问题检出率大小不一的关键因素。(3)除了睡眠问题外，检出时间对小学生心理健康指标检出率的影响较大。(4)虽然区域对攻击行为指标有显著影响，但整体而言，区域对小学生心理健康问题检出率的影响不大。



## 参考文献

陈洁琼. (2012). *儿童焦虑情绪症状的影响因素及干预研究*(硕士学位论文). 苏州大学.

程琳. (2010). *父母期望、初中生自我期望与学习成绩的关系*(硕士学位论文). 河南大学.

杜柏玲. (2016). 甘肃山区农村小学生心理健康状况及其影响因素研究——以陇南市为例. *甘肃高师学*报, *21*(11), 54–58.

傅小兰, 张侃, 陈雪峰, 陈祉妍. (2021). 心理健康蓝皮书: 中国国民心理健康发展报告(2019-2020). 北京: 社会科学文献出版社.

黄文治, 王志刚. (2020). 小学生学习压力与心理健康的关系研究. *校园心理, 18*(1), 36–38.

来枭雄, 黄顺森, 张彩, 唐斌, 张美萱, 朱成伟, 王耘. (2020). 中小学生手机成瘾与人际关系主观幸福感和学校认同感的关联. *中国学校卫生, 41*(4), 613–616.

李超群, 徐良苑, 李燕芳. (2021). 教师性别刻板印象对中小学生性别刻板印象的影响：教师性别教育行为的中介作用. *中国特殊教育*, *4*, 90–96.

李梦龙, 任玉嘉, 蒋芬. (2019). 中国农村留守儿童社交焦虑状况的 Meta 分析. *中国心理卫生杂志, 33*(11), 839–844.

廖友国, 连榕. (2019). 近三十年国民心理健康变迁的横断历史研究. *西南大学学报(社会科学版), 45*(2), 105–116.

刘福荣, 吴梦凡, 董一超, 岳馨培, 尚小平, 隋美丽, 刘新奎. (2021). 小学生抑郁症状检出率的 meta 分析. *中国心理卫生杂志, 35*(6), 482–488.

刘国雄, 张丽锦. (2010). 关于情绪以及情绪发展的理论述评. *宁夏大学学报(人文社会科学版), 32*(1), 212–215.

刘宇, 胡传鹏, 樊富珉, 孙沛, 徐杰, 蔡玉清等. 基于网络理论的物质成瘾新视角. *心理科学进展, 29*(2), 296–310.

龙鑫, 纪颖, 张洪伟, 张夏男, 谢立璟. (2020). 北京市中高年级小学生睡眠时间现状及影响因素分析. *中国卫生统计, 37*(5), 704–707.

王燕, 刘思洁, 陈矜之. (2020). 改革开放 40 年中国人心态变化的年代分析——以社会信任, 主观幸福感和心理健康为例. *苏州大学学报(教育科学版), 8*(1), 58–69.

卫萍, 葛明贵, 陈雪梅. (2007). 合肥848名小学生心理健康状况的调查分析. *合肥学院学报: 社会科学版, 24*(06), 40–43.
忻仁娥, 张志雄. (1992). 全国22个省市26个单位24013名城市在校少年儿童行为问题调查:—独生子女精神卫生问题的调查,防治和Achenbach's儿童行为量表中国标准化. *上海精神医学, 4*(1), 47–55.
辛自强, 张梅. (2009). 1992年以来中学生心理健康的变迁:一项横断历史研究. *心理学报, 41*(1), 69–78.
徐美贞. (2005). 中小学心理健康教育现状调查分析. *教育发展研究, 6*, 78–83.
杨洪猛, 高松, 傅金芝. (2003). 少数民族中、小学生心理健康状况研究. *中国健康心理学杂志, 11*(6), 416–418.
俞国良, 李天然, 王勍. (2016). 高中生心理健康的横断历史研究. *教育研究, 37*(10), 113–122.
俞国良, 赵凤青, 罗晓路. (2017). 心理健康教育:高等院校的地区差异比较研究. *黑龙江高教研究, 12*, 89–92.
余欣欣, 王洁莹, 杨静. (2019). 广西三～六年级小学生心理健康现状分析. *中国健康教育, 35*(12), 1089–1093.
张斌, 郝彦利, 任衍镇. (2013). 广州市小学生的学业成绩与睡眠状况. *中国心理卫生杂志, 27*(6), 462–467.
张柳, 荆玉梅. (2019). 黄石市316名小学生心理健康状况分析及对策研究. *湖北师范大学学报: 哲学社会科学版, 39*(4), 112–116.
赵建华. (2005). 小学生攻击性行为的心理分析及对策研究. *心理科学, 28*(4), 965–968.
赵玫. (2010). 农村小学生心理健康状况分析. *卫生职业教育, 28*(9), 112–113.
Achenbach TM. (1966). The classification of children's psychiatric symptoms: A factor-analytic study. Psychol Monogr, *80*, 1–37.
Achenbach, T. M., & Edelbrock, C. S. (1979). The child behavior profile: II. boys aged 12-16 and girls aged 6-11 and 12-16. *Journal of Consulting and Clinical Psychology, 47*(2), 223–233.
Birleson, P. (1981). The validity of depressive disorder in childhood and the development of a self-rating scale: A research report. *Journal of Child Psychology and Psychiatry, 22*(1), 73–88.
Birmaher, B., Khetarpal, S., Brent, D., Cully, M., Balach, L., Kaufman, J., & Neer, S. M. (1997). The screen for child anxiety related emotional disorders (SCARED): Scale construction and psychometric characteristics. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*(4), 545–553.
Buss, A. H., & Warren, W. L. (2000). *The aggression questionnaire manual*. Los Angeles: Western Psychological Services.
Buysse, D. J., Iii, C., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28*(2), 193–213.
Card, N. A. (2012). *Applied meta-analysis for social science research*. New York: Guilford Press.
Chen, X., ling Ke, Z., Chen, Y., & Lin, X. (2021). The prevalence of sleep problems among children in mainland China: A meta-analysis and systemic-analysis. *Sleep Medicine*, *83*, 248–255.
Conners, C. K. (1969). A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry*, *126*(6), 884–888.
Erskine, H. E., Baxter, A. J., Patton, G., Moffitt, T. E., Patel, V., Whiteford, H. A., & Scott, J. G. (2017). The global coverage of prevalence data for mental disorders in children and adolescents. *Epidemiology and Psychiatric Sciences*, *26*(4), 395–402.
Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*, 557–560.
Jaafari, Z., Farhadi, A., Amin Lari, F., Mousavi, F. S., Moltafet, H., Dashti, E., & Marzban, M. (in press). Prevalence of Depression in Iranian College Students: A Systematic Review and Meta-analysis. *Iranian Journal of Psychiatry and Behavioral Sciences*.
Kovacs, M. (1992). *Children's Depression Inventory (CDI)*. New York: Multi-health Systems, Inc.
Kuppens, S., Laurent, L., Heyvaert, M., & Onghena, P. (2013). Associations between parental psychological control and relational aggression in children and adolescents: A multilevel and sequential meta-

analysis. *Developmental Psychology*, *49*(9), 1697.

Li, J. Y., Li, J., Liang, J. H., Qian, S., Jia, R. X., Wang, Y. Q., & Xu, Y. (2019). Depressive symptoms among children and adolescents in China: A systematic review and meta-analysis. *Medical Science Monitor, 25,* 74–59.

Owens, J. A., Anthony, S., & Melissa, M. G. (2000). The children's sleep habits questionnaire (CSHQ): Psychometric properties of a survey instrument for school-aged children. *Sleep, 23*(8), 1043–1051.

Radloff, L. S. (1977). A self-report depression scale for research in the general population. *Applied Psychol Measurements, 1*(3), 385–401.

Wood, A., Kroll, L., Moore, A., & Harrington, R. (1995). Properties of the mood and feelings questionnaire in adolescent psychiatric outpatients: A research note. *Journal of Child Psychology and Psychiatry*, *36*(2), 327–334.



# Prevalence of mental health problems among primary school students in China from 2010 to 2010: A meta-analysis

HUANG Xiaoxiao[1], ZHANG Yali[1], YU Guoliang[2]

([1] *School of Education, Renmin University of China, Beijing* 100872, *China*)

([2] *Institute of Psychology, Renmin University of China, Beijing* 100872, *China*)

**Abstract:** The prevalence of mental health problems of primary school students in China were different. In order to integrate the survey results and analyze the causes, meta-analysis was used to synthesize the relevant studies in recent ten years from 2010 to 2020 (101 articles, 289396 subjects). The results showed that: (1) The prevalence of mental health problems in primary school students from high to low were sleep problems (25.2%, 95%CI = [0.160, 0.373]), depression (14.6%, 95%CI = [0.120, 0.175]), anxiety (12.3%, 95%CI = [0.063, 0.227]), aggressive behavior (4.1%, 95%CI = [0.017, 0.099]), withdrawal (3.8%, 95%CI = [0.023, 0.062]), disciplinary behavior (3.7%, 95%CI = [0.018, 0.073]) and somatization (3.6%, 95%CI = [0.017, 0.073]). (2) Measurement tools, measurement standards and detection period were the key factors causing the fluctuation of prevalence of mental health problems among primary school students in China. Generally speaking, the overall mental health status of primary school students in China is acceptable except that the prevalence of sleep problems, depression and anxiety are high. In the future, we should develop mental health assessment tools and detection standards suitable for Chinese primary school students to provide support for the prevention and accurate intervention of mental health problems.

**Key words:** pupils; mental health problems; prevalence; meta-analysis